

POWER-LAW DECAY OF THE DEGREE-SEQUENCE PROBABILITIES OF TWO RANDOM GRAPHS WITH APPLICATION TO GRAPH ISOMORPHISM

JEFFERSON ELBERT SIMÕES

DANIEL R. FIGUEIREDO

VALMIR C. BARBOSA

*Systems Engineering and Computer Science Program, COPPE
Federal University of Rio de Janeiro, Rio de Janeiro, Brazil*

ABSTRACT. We consider events over the probability space generated by the degree sequences of two independent Erdős-Rényi random graphs, and consider an approximation probability space where such degree sequences are deemed to be sequences of i.i.d. random variables. We show that, for any sequence of events with probabilities asymptotically smaller than some power law in the approximation model, the same upper bound also holds in the original model. We accomplish this by extending an approximation framework proposed in a seminal paper by McKay and Wormald. Finally, as an example, we apply the developed framework to bound the probability of having an isomorphism between two independent random graphs.

1. INTRODUCTION

The Erdős-Rényi random graph model, also known as the $G(n, p)$ model [7][9] is the most traditional probabilistic model for graphs. Introduced in 1959, in this model, a graph over n vertices is randomly generated by adding edges independently between each vertex pair with probability $p(n)$. Despite its inability to model real-world networks, its simplicity and the consequent analytical tractability have allowed thorough theoretical analysis [5] and applications such as percolation models [15] and graph theory via the probabilistic method [1].

One of the toughest challenges in understanding the overall structure of the $G(n, p)$ random graph is obtaining a precise characterization of its degree sequence. The main reason for this is that, even though the degrees of any two specific nodes are only mildly correlated (due to the possible edge between them), it is still a nontrivial task to compose these correlations into a manageable joint distribution for the degrees.

Most results on this matter address the distribution of the t -th largest degree, for some $t(n)$ generally bounded. More recently, though, a framework has been set by McKay and Wormald [13] for approximating the degree sequence by a sequence of independent random variables, with tight bounds on the error of the probabilities

E-mail addresses: {elbert,daniel}@land.ufrj.br, valmir@cos.ufrj.br.

This work has been partially funded by CAPES, CNPq, and FAPERJ BBP grants.

of events estimated by this approximation. This framework has been successfully applied in several contexts: for instance, Kostochka and West [11] use it to analyze the middle degree asymptotics of random graphs, which relates to Chvátal's condition for Hamiltonian graphs, and Skerman [16] applies a similar technique to analyze degrees in a random bipartite graph model.

In this paper, we consider the problems of comparing the degree sequences of two random graphs, and of approximating these degree sequences by corresponding sequences of independent random variables. Our main result (Theorem 3.2) directly relates power-law decaying probabilities in the two models: any event sequence that has probability $o(n^{-a})$ in the approximation model also has probability $o(n^{-a})$ in the original degree sequence model. To achieve this, we extend the framework in [13] to establish a relationship between the degree sequences of both graphs and the corresponding independent sequences through a series of intermediate approximations. The stepwise error bounds, formally established by Theorem 3.1, lay down a roadmap for handling asymptotic probabilities of properties that compare the structures of a pair of graphs.

As an example, we apply Theorem 3.2 to the problem of graph isomorphism. Not only is this problem an interesting theoretical problem in its own right, but it also has implications in practical problems such as network privacy and anonymization [14] and computer vision [8]. In particular, we show that, for a certain range of model parameters, two random graphs will not be isomorphic with probability $1 - o(n^{-1/2})$.

This paper is structured as follows: in section 2 we review the degree sequence approximation framework, detailing its steps and stating the main results used. We then proceed to extending the framework to two independent random graphs, providing corresponding statements and proofs in section 3. Our sample application will be presented in section 4, where we apply the framework to the problem of isomorphism, after which we conclude with some final remarks in section 5.

2. RELATED WORK

McKay and Wormald [13] have previously formalized, under quite loose constraints, the very intuitive result that the degree sequence of a $G(n, p)$ random graph is similar to a sequence of independent random variables, each having distribution $\text{Bin}(n-1, p)$. This result takes the form of a number of theorems and lemmas, each performing one of four steps in the approximation process that is detailed in this section. Notation will be kept as similar as possible to the original work [13].

For some fixed $n \in \mathbb{N}$, take the set $I_n = \{0, \dots, n-1\}^n$ equipped with the discrete σ -algebra as our measurable space. Let $d = (d_1, \dots, d_n)$ be some element in this space. Also, let $p = p(n) \in (0, 1)$, and denote $N = \binom{n}{2}$ and $q = 1 - p$.

In the *binomial model* $\mathcal{B}_{n,p}$, d is distributed as a sequence of n independent $\text{Bin}(n-1, p)$ random variables. This can be achieved by evaluating d under the probability measure $\mathbb{P}_{\mathcal{B}_{n,p}} = \text{Bin}(n-1, p)^{\otimes n}$. We would like to assert that this model is similar to the degree sequence of a $G(n, p)$ random graph. We call this the *degree sequence model* $(\mathcal{D}_{n,p})$, and denote by $\mathbb{P}_{\mathcal{D}_{n,p}}$ the probability measure under which d has this distribution. Note that the sum of degrees in any graph is necessarily even, which means d will take, with probability 1, values on the set $E_n = \{d \in I_n : M(d) \text{ is even}\}$ (where $M = M(d) = \|d\|_1$ is the sum of the components of d).

The approximation process requires three additional models (with corresponding probability measures) that will perform a transition from the binomial model to the degree sequence model, with two of them making d acquire properties from the

degree sequence model that are not present in the binomial model, and the third one acting as a technical middleman. The first model is the *even-sum binomial model* ($\mathcal{E}_{n,p}$). It ensures that d indeed takes values in E_n with probability 1. To ensure minimum distortion between probability of elements of E_n , this model is simply set to be the restriction of the binomial model to the set E_n .¹ Then, the *weighted even-sum binomial model* ($\mathcal{E}'_{n,p}$) ensures the stronger property that M has the same distribution as it does under the degree sequence model (namely, that $M/2$ is distributed as $\text{Bin}(N, p)$). To insert as little interference as possible into the relative probabilities of any two points in E_n , the probabilities of all points E_n are rescaled (or *reweighted*) uniformly on each set $S_m = \{d \in E_n : M(d) = m\}$, to make these sets have the desired probability.

To perform the bridge between $\mathcal{E}'_{n,p}$ and $\mathcal{E}_{n,p}$, they have introduced the *integrated model* $\mathcal{I}_{n,p}$, which is essentially a “noisy” version of the even-sum model $\mathcal{E}_{n,p}$. The model $\mathcal{I}_{n,p}$ is obtained from $\mathcal{E}_{n,p}$ by switching from a fixed parameter p to a random parameter p' that quickly concentrates around p . More specifically, p' must be distributed as a truncated normal variable, with expected value p , variance $pq/2N$, and restricted to the unit interval.

We can informally summarize the approximation scheme as follows:

$$\mathbb{P}_{\mathcal{B}_{n,p}} \approx \mathbb{P}_{\mathcal{E}_{n,p}} \approx \mathbb{P}_{\mathcal{I}_{n,p}} \approx \mathbb{P}_{\mathcal{E}'_{n,p}} \approx \mathbb{P}_{\mathcal{D}_{n,p}}$$

Now, for these approximations to work, it is necessary for $p(n)$ to lie in a “good behavior range”, in which case $p = p(n)$ is said to be *acceptable*. The last approximation, in particular, is hard to tighten in general, so the necessary conditions for this approximation to work are brought into the definition of acceptable function:

Definition 2.1. Let $\lambda = \lambda(d) = M(d)/2N$ and $\gamma_2 = \gamma_2(d) = (n-1)^{-2} \sum_{i=1}^n (d_i - M(d))^2$. A function $p = p(n)$ is acceptable if the following conditions hold:

- (1) $pqN = \omega(n) \log n$;
- (2) there is a set $R_p(n) \subset E_n$ and a real function $\delta(n) = o(1)$ such that:
 - (a) $\mathbb{P}_{\mathcal{D}_{n,p}}(R_p(n)), \mathbb{P}_{\mathcal{E}_{n,p}}(R_p(n)) = 1 - n^{-\omega(n)}$;
 - (b) for every $d \in R_p(n)$, there is some δ_d such that $|\delta_d| \leq \delta(n)$ and

$$\frac{\mathbb{P}_{\mathcal{D}_{n,p}}(d)}{\mathbb{P}_{\mathcal{E}'_{n,p}}(d)} = \exp \left\{ \frac{1}{4} \left(1 - \frac{\gamma_2^2}{\lambda^2(1-\lambda)^2} \right) \right\} \cdot \exp\{\delta_d\}.$$

The second condition in this definition requires a set $R_p(n)$ to exist in our sample space E_n , with very large probability in $\mathcal{D}_{n,p}$ and $\mathcal{E}_{n,p}$ (the probability of its complement in both models vanishes faster than any standard exponential), in which the models $\mathcal{D}_{n,p}$ and $\mathcal{E}'_{n,p}$ uniformly agree to a ratio that approaches 1. This condition is required for the proofs to be carried out, though it has been conjectured by McKay and Wormald that condition 1 in the definition is sufficient for $p(n)$ to be acceptable — to the best of our knowledge, this conjecture is still open. For our purposes, they have identified an interesting regime for $p(n)$ in which these conditions hold [13]:

Theorem 2.2. $p(n)$ is acceptable whenever $\omega(n) \log n/n^2 \leq pq \leq o(n^{-1/2})$.

The execution of this approximation scheme has been broken down into a number of pieces with various levels of complexity, so to fit different possibilities of applications. In our particular case, we would like to ensure that this scheme is well-suited for approximating probabilities that vanish faster than power laws in n . For this purpose, we extract the following results from [13], condensed in a single theorem.

¹ That is, the corresponding probability measure is the measure for the binomial model conditional to the event E_n , evaluated only on the events E_n .

Theorem 2.3. Write $\phi(x; \mu, \sigma^2)$ the density function of the normal distribution, and $V_{n,p} = \int_0^1 \phi(x; p, pq/2N) dx$. Then the following statements hold:

(1) For any event $A_n \subseteq E_n$,

$$\mathbb{P}_{\mathcal{E}_{n,p}}(A_n) = \frac{2\mathbb{P}_{\mathcal{B}_{n,p}}(A_n)}{1 + (q-p)^{2N}};$$

(2) For any event $A_n \subseteq E_n$,

$$\mathbb{P}_{\mathcal{I}_{n,p}}(A_n) = \frac{1}{V_{n,p}} \int_0^1 \phi(x; p, pq/2N) \mathbb{P}_{\mathcal{E}_{n,x}}(A_n) dx;$$

(3) If $pqN \rightarrow \infty$ and $y = y(n) = o(\sqrt[6]{pqN})$, then

$$\mathbb{P}_{\mathcal{I}_{n,p}}(d) = \mathbb{P}_{\mathcal{E}'_{n,p}}(d) \left(1 + O\left(\frac{1 + |y|^3}{\sqrt{pqN}}\right) \right)$$

uniformly over $\{d \in E_n : |M(d) - 2Np| \leq 2y\sqrt{Npq}\}$;

(4) If $\omega(n) \log n/n^2 \leq pq \leq o(n^{-1/2})$, then there are sets $R_p(n), R'_p(n) \subseteq E_n$ and a real function $\delta(n) = o(1)$ such that:

(a) $\mathbb{P}_{\mathcal{D}_{n,p}}(R_p(n)), \mathbb{P}_{\mathcal{D}_{n,p}}(R'_p(n)) = 1 - n^{-\omega(n)}$;

(b) in $R'_p(n)$, $\gamma_2 = \lambda(1 - \lambda)(1 + o(1))$;

(c) for every $d \in R_p(n)$, there is some δ_d such that $|\delta_d| \leq \delta(n)$ and

$$\frac{\mathbb{P}_{\mathcal{D}_{n,p}}(d)}{\mathbb{P}_{\mathcal{E}'_{n,p}}(d)} = \exp \left\{ \frac{1}{4} \left(1 - \frac{\gamma_2^2}{\lambda^2(1 - \lambda)^2} \right) \right\} \cdot \exp\{\delta_d\}.$$

Proof. All referenced results used in this proof have been extracted from [13], to which we refer the reader for notation and statements. Statement 1 is a particular case of corollary 4.3 taking $f = \mathbb{I}_{A_n}$ the indicator function of the event A_n , simplified by theorem 4.2 and the observation that, since $f = 0$ in $I_n \setminus E_n$, $f = \tilde{f}$. Statement 2 is a rewriting of lemma 2.4, consequence of the construction of $\mathbb{P}_{\mathcal{I}_{n,p}}$ from $\mathbb{P}_{\mathcal{E}_{n,p}}$ and an application of the law of total probability — we note that, for $x \in [0, 1]$, $\phi(x; p, pq/2N)/V_{n,p}$ is the density function of the random parameter p' used in the construction. Statement 3 simply restates theorem 3.6. Statement 4 comes from the definition of acceptability and corollary 3.5, noting that the hypothesis implies $p(n)$ is acceptable. \square

Theorem 2.3 suffices for us to ensure suitability of the scheme to our purposes:

Theorem 2.4. Let A_n be a sequence of events in E_n . If p satisfies $\omega(\log n/n) \leq p \leq o(n^{-1/2})$, then for any fixed $a > 0$, $\mathbb{P}_{\mathcal{B}_{n,p}}(A_n) = o(n^{-a})$ implies $\mathbb{P}_{\mathcal{D}_{n,p}}(A_n) = o(n^{-a})$.

Each step in the proof of this theorem is a simplified version of the corresponding step in the proof of Theorem 3.2, which considers two independent random graphs, to be presented in the next section. For brevity, we will explicitly provide proof only for the latter theorem.

It is worth noting that, in the abstract of their paper, McKay and Wormald state that their techniques can be used to determine highly accurate asymptotics for probabilities that, in $\mathcal{D}_{n,p}$, are greater than any (fixed) power law. While this statement, and particularly the meaning of “highly accurate”, has been posed in an informal fashion, it can be understood, in light of their approximation scheme, to mean the following formal statement:

If $p(n)$ is acceptable, then for any event $A_n \subseteq E_n$, if $\mathbb{P}_{\mathcal{D}_{n,p}}(A_n) = \omega(n^{-k})$ for some fixed k , then there are functions $\alpha(n) = \theta(1)$ and $\varepsilon(n) = n^{-\omega(n)}$ such that $\mathbb{P}_{\mathcal{D}_{n,p}}(A_n) = \mathbb{P}_{\mathcal{B}_{n,p}}(A_n) \cdot \alpha(n) + \varepsilon(n)$.

Our result could, in principle, be derived as a corollary of this fact since it asserts that, if $p(n)$ is acceptable, $\mathbb{P}_{\mathcal{D}_{n,p}}(A_n) = \Omega(n^{-k})$ implies $\mathbb{P}_{\mathcal{B}_{n,p}}(A_n) = \Omega(n^{-k})$, which is the contrapositive of our result. However, since this fact has neither been formally stated nor proven, we choose to prove our result directly by using the same fundamental techniques (i.e., individual pieces of the approximation scheme) that would be required to prove the aforementioned fact.

3. RESULTS

In several domains, we can identify problems that can be reduced to understanding whether the structure of two given graphs are similar. In this work, we consider the situation where these two graphs are given instances of the $G(n, p)$ model, with the same size but possibly with different values of p — that is, two random graphs $G(n, p)$ and $G(n, p')$. We also assume that these instances are independent.

Naturally, we would like to compare the degree sequences of these two graphs, as such comparison can be used as a proxy for more complicated properties. Intuitively, it would seem trivial that, since the two degree sequences are independent of each other and can both be individually approximated by i.i.d. sequences with small errors on the corresponding probabilities of events, the joint approximation of both degree sequences should similarly yield a small error as well. However, we find it essential that this extension of the single-graph case be obtained formally. As we see in what follows, even though such extension is indeed possible, achieving it is far from trivial.

Before we proceed, let us introduce some notation. For $p, p' \in (0, 1)$, denote by $\mathbb{P}_{\mathcal{B}_{n,p,p'}}$ the probability measure $\mathbb{P}_{\mathcal{B}_{n,p}} \otimes \mathbb{P}_{\mathcal{B}_{n,p'}}$ over I_n^2 — and similarly for measures in other models, over E_n^2 . Our goal is to perform the following approximation scheme:

$$\mathbb{P}_{\mathcal{B}_{n,p,p'}} \approx \mathbb{P}_{\mathcal{E}_{n,p,p'}} \approx \mathbb{P}_{\mathcal{I}_{n,p,p'}} \approx \mathbb{P}_{\mathcal{E}'_{n,p,p'}} \approx \mathbb{P}_{\mathcal{D}_{n,p,p'}}$$

Let us stress that $\mathbb{P}_{\mathcal{D}_{n,p,p'}} = \mathbb{P}_{\mathcal{D}_{n,p}} \otimes \mathbb{P}_{\mathcal{D}_{n,p'}}$ is the distribution of the degree sequence of two random graphs $G(n, p)$ and $G(n, p')$ independent of each other, and $\mathbb{P}_{\mathcal{B}_{n,p,p'}}$ is the corresponding approximation by two independent sequences of i.i.d. random variables.

We will extend our notation further and write $q' = 1 - p'$, and denote by $d \times d'$ some element of I_n^2 . We will also write $\lambda = \lambda(d \times d') = M(d)/2N$, $\lambda' = \lambda'(d \times d') = M(d')/2N$, $\gamma_2 = \gamma_2(d \times d') = (n-1)^{-2} \sum_{i=1}^n (d_i - M(d))^2$ and $\gamma'_2 = \gamma'_2(d \times d') = (n-1)^{-2} \sum_{i=1}^n (d'_i - M(d'))^2$.

In the following, we will prove that an extended version of Theorem 2.3 holds for these models:

Theorem 3.1. *Let $\phi(x; \mu, \sigma^2)$ and $V_{n,p}$ as in Theorem 2.3. Then the following statements hold:*

- (1) *For any event $A_n \subseteq E_n^2$,*

$$\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) = \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n)}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]};$$

- (2) *For any event $A_n \subseteq E_n^2$,*

$$\mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n) = \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) dx dx';$$

- (3) *If $pqN, p'q'N \rightarrow \infty$ and $y = y(n)$ is both $o(\sqrt[6]{pqN})$ and $o(\sqrt[6]{p'q'N})$, then*

$$\mathbb{P}_{\mathcal{I}_{n,p,p'}}(d \times d') = \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d') \left(1 + O\left(\frac{1 + |y|^3}{\sqrt{pqN}}\right) + O\left(\frac{1 + |y|^3}{\sqrt{p'q'N}}\right) \right)$$

- uniformly over $\{d \times d' \in E_n^2 : |M(d) - 2Np| \leq 2y\sqrt{Npq}, |M(d') - 2Np'| \leq 2y\sqrt{Np'q'}\}$;
- (4) If $\omega(n) \log n/n^2 \leq pq, p'q' \leq o(n^{-1/2})$, then there are sets $S_{p,p'}(n) \subseteq E_n^2$ and $S'_{p,p'}(n) \subseteq E_n^2$ and a real function $\varepsilon(n) = o(1)$ such that:
- (a) $\mathbb{P}_{\mathcal{D}_{n,p,p'}}(S_{p,p'}(n)), \mathbb{P}_{\mathcal{D}_{n,p,p'}}(S'_{p,p'}(n)) = 1 - n^{-\omega(n)}$;
 - (b) in $S'_{p,p'}(n)$, $\gamma_2 = \lambda(1 - \lambda)(1 + o(1))$ and $\gamma'_2 = \lambda'(1 - \lambda')(1 + o(1))$;
 - (c) for every $d \times d' \in S_{p,p'}(n)$, there is some $\varepsilon_{d \times d'}$ such that $|\varepsilon_{d \times d'}| \leq \varepsilon(n)$ and

$$\frac{\mathbb{P}_{\mathcal{D}_{n,p,p'}}(d \times d')}{\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d')} = \exp \left\{ \frac{1}{4} \left(2 - \frac{\gamma_2^2}{\lambda^2(1 - \lambda)^2} - \frac{(\gamma'_2)^2}{(\lambda')^2(1 - \lambda')^2} \right) \right\} \cdot \exp\{\varepsilon_{d \times d'}\}.$$

Proof. **Statement 1:** Let \mathcal{F} be the family of subsets of E_n^2 for which the statement's equality holds. We will prove that (i) \mathcal{F} contains all rectangles (i.e., events of the form $B \times B'$, with $B, B' \subset E_n$) and (ii) \mathcal{F} is a λ -system. This is enough since, by Dynkin's theorem, \mathcal{F} must contain the σ -algebra generated by the rectangles, which is the discrete σ -algebra over E_n^2 .

For the first claim, for any rectangle $A = B \times B'$ in E_n^2 , by Theorem 2.3(1), we have that

$$\begin{aligned} \mathbb{P}_{\mathcal{E}_{n,p,p'}}(A) &= \mathbb{P}_{\mathcal{E}_{n,p}} \otimes \mathbb{P}_{\mathcal{E}_{n,p'}}(B \times B') \\ &= \mathbb{P}_{\mathcal{E}_{n,p}}(B) \mathbb{P}_{\mathcal{E}_{n,p'}}(B') \\ &= \frac{2\mathbb{P}_{\mathcal{B}_{n,p}}(B)}{1 + (q - p)^{2N}} \cdot \frac{2\mathbb{P}_{\mathcal{B}_{n,p'}}(B')}{1 + (q' - p')^{2N}} \\ &= \frac{4\mathbb{P}_{\mathcal{B}_{n,p}}(B) \mathbb{P}_{\mathcal{B}_{n,p'}}(B')}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]} \\ &= \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A)}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]}. \end{aligned}$$

Therefore, \mathcal{F} contains all rectangles.

For the second claim, note that \mathcal{F} contains E_n^2 , since it is a rectangle; \mathcal{F} is closed by complements, since for any $A \in \mathcal{F}$, it holds that

$$\begin{aligned} \mathbb{P}_{\mathcal{E}_{n,p,p'}}(E_n^2 \setminus A) &= \mathbb{P}_{\mathcal{E}_{n,p,p'}}(E_n^2) - \mathbb{P}_{\mathcal{E}_{n,p,p'}}(A) \\ &= \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(E_n^2)}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]} - \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A)}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]} \\ &= \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(E_n^2 \setminus A)}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]} \end{aligned}$$

and $E_n^2 \setminus A \in \mathcal{F}$; and \mathcal{F} is also closed by disjoint enumerable unions, since for any sequence B_1, B_2, \dots in \mathcal{F} , if B_1, B_2, \dots are disjoint, then

$$\begin{aligned}
\mathbb{P}_{\mathcal{E}_{n,p,p'}} \left(\biguplus_{i=1}^{\infty} B_i \right) &= \sum_{i=1}^{\infty} \mathbb{P}_{\mathcal{E}_{n,p,p'}}(B_i) \\
&= \sum_{i=1}^{\infty} \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(B_i)}{[1 + (q-p)^{2N}][1 + (q'-p')^{2N}]} \\
&= \frac{4 \sum_{i=1}^{\infty} \mathbb{P}_{\mathcal{B}_{n,p,p'}}(B_i)}{[1 + (q-p)^{2N}][1 + (q'-p')^{2N}]} \\
&= \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(\biguplus_{i=1}^{\infty} B_i)}{[1 + (q-p)^{2N}][1 + (q'-p')^{2N}]}
\end{aligned}$$

and $\biguplus_{i=1}^{\infty} B_i \in \mathcal{F}$. Since \mathcal{F} fits the three requirements, by definition, \mathcal{F} is a λ -system.

Statement 2: We will use the same strategy as in statement 1. Let \mathcal{G} be the family of subsets of E_n^2 for which the statement's equality is true. First, take an arbitrary rectangle $A = B \times B'$ in E_n^2 . Using Theorem 2.3(2) yields

$$\begin{aligned}
\mathbb{P}_{\mathcal{I}_{n,p,p'}}(A) &= \mathbb{P}_{\mathcal{I}_{n,p}} \otimes \mathbb{P}_{\mathcal{I}_{n,p'}}(B \times B') \\
&= \mathbb{P}_{\mathcal{I}_{n,p}}(B) \mathbb{P}_{\mathcal{I}_{n,p'}}(B') \\
&= \frac{1}{V_{n,p}} \int_0^1 \phi(x; p, pq/2N) \mathbb{P}_{\mathcal{E}_{n,x}}(B) dx \\
&\quad \cdot \frac{1}{V_{n,p'}} \int_0^1 \phi(x'; p', p'q'/2N) \mathbb{P}_{\mathcal{E}_{n,x'}}(B') dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,p}}(B) \mathbb{P}_{\mathcal{E}_{n,p'}}(B') dx dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(B) dx dx',
\end{aligned}$$

which means \mathcal{G} contains all rectangles.

Secondly, \mathcal{G} satisfies the three requirements of the definition of λ -systems: it contains E_n^2 , since it is a rectangle; it is closed under complements, since

for any $A \in \mathcal{G}$,

$$\begin{aligned}
\mathbb{P}_{\mathcal{I}_{n,p,p'}}(E_n^2 \setminus A) &= \mathbb{P}_{\mathcal{I}_{n,p,p'}}(E_n^2) - \mathbb{P}_{\mathcal{I}_{n,p,p'}}(A) \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(E_n^2) dx dx' \\
&\quad - \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A) dx dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \left[\phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(E_n^2) \right. \\
&\quad \left. - \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A) \right] dx dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \\
&\quad \cdot [\mathbb{P}_{\mathcal{E}_{n,x,x'}}(E_n^2) - \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A)] dx dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(E_n^2 \setminus A) dx dx'
\end{aligned}$$

and $E_n^2 \setminus A \in \mathcal{G}$; and \mathcal{F} is also closed by disjoint enumerable unions, since for any sequence B_1, B_2, \dots in \mathcal{G} , if B_1, B_2, \dots are disjoint, then

$$\begin{aligned}
&\mathbb{P}_{\mathcal{I}_{n,p,p'}}\left(\biguplus_{i=1}^{\infty} B_i\right) \\
&= \sum_{i=1}^{\infty} \mathbb{P}_{\mathcal{I}_{n,p,p'}}(B_i) \\
&= \sum_{i=1}^{\infty} \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(B_i) dx dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \left[\sum_{i=1}^{\infty} \mathbb{P}_{\mathcal{E}_{n,x,x'}}(B_i) \right] dx dx' \\
&= \frac{1}{V_{n,p}V_{n,p'}} \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(\biguplus_{i=1}^{\infty} B_i) dx dx'
\end{aligned}$$

and $\biguplus_{i=1}^{\infty} B_i \in \mathcal{G}$. Since \mathcal{G} is a λ -system and contains all rectangles, by Dynkin's theorem, it must also contain the σ -algebra generated by the rectangles, which is the discrete σ -algebra over E_n .

Statement 3: Take $d \times d' \in E_n^2$ satisfying both $|M(d) - 2Np| \leq 2y\sqrt{Npq}$ and $|M(d') - 2Np'| \leq 2y\sqrt{Np'q'}$. Then, using Theorem 2.3(3) we can write

$$\begin{aligned}
&\mathbb{P}_{\mathcal{I}_{n,p,p'}}(d \times d') \\
&= \mathbb{P}_{\mathcal{I}_{n,p}}(d) \cdot \mathbb{P}_{\mathcal{I}_{n,p'}}(d') \\
&= \mathbb{P}_{\mathcal{E}'_{n,p}}(d) \left(1 + O\left(\frac{1+|y|^3}{\sqrt{pqN}}\right) \right) \mathbb{P}_{\mathcal{E}'_{n,p'}}(d') \left(1 + O\left(\frac{1+|y|^3}{\sqrt{p'q'N}}\right) \right).
\end{aligned}$$

Note that the inequality from Theorem 2.3(3) was applied twice, for $\mathcal{I}_{n,p}$ and for $\mathcal{I}_{n,p'}$. Since both inequalities are uniform in their respective sets — $\{d \in E_n : |M(d) - 2Np| \leq 2y\sqrt{Npq}\}$ and $\{d' \in E_n : |M(d') - 2Np'| \leq 2y\sqrt{Np'q'}\}$ —, the resulting inequality is uniform in the set $\{d \times d' \in$

$E_n^2 : |M(d) - 2Np| \leq 2y\sqrt{Npq}, |M(d') - 2Np'| \leq 2y\sqrt{Np'q'}\}$. Algebraic manipulations yield

$$\begin{aligned} \mathbb{P}_{\mathcal{I}_{n,p,p'}}(d \times d') &= \mathbb{P}_{\mathcal{E}'_{n,p}}(d) \mathbb{P}_{\mathcal{E}'_{n,p'}}(d') \left(1 + O\left(\frac{1+|y|^3}{\sqrt{pqN}}\right) + O\left(\frac{1+|y'|^3}{\sqrt{p'q'N}}\right) \right. \\ &\quad \left. + O\left(\frac{1+|y|^3}{\sqrt{pqN}}\right) O\left(\frac{1+|y'|^3}{\sqrt{p'q'N}}\right) \right) \\ &= \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d') \left(1 + O\left(\frac{1+|y|^3}{\sqrt{pqN}}\right) + O\left(\frac{1+|y'|^3}{\sqrt{p'q'N}}\right) \right), \end{aligned}$$

since $y = o(\sqrt[3]{pqN})$ (which implies $(1+|y|^3)/\sqrt{pqN} = o(1)$). This proves the result.

Statement 4: This proof will follow by construction. Under the stated assumptions for p , there exist sets $R_p(n), R'_p(n) \subseteq E_n$ and a real function $\delta(n)$ satisfying the conditions of Theorem 2.3(4), with δ_d for each $d \in R_p(n)$ in condition (b). Applying the same reasoning for p' , there are also sets $R_{p'}(n), R'_{p'}(n) \subseteq E_n$ and real function $\delta'(n)$ satisfying these same conditions, this time with $\delta'_{d'}$ for each $d' \in R_{p'}(n)$ in condition (b) (δ and δ' will not necessarily be equal for $d = d'$). Note that $\delta(n), \delta'(n)$ must be positive real functions.

Now, take:

$$\begin{aligned} S_{p,p'}(n) &= R_p(n) \times R_{p'}(n), \\ S'_{p,p'}(n) &= R'_p(n) \times R'_{p'}(n), \\ \varepsilon(n) &= \delta(n) + \delta'(n). \end{aligned}$$

We will show the desired results hold for $S_{p,p'}$, $S'_{p,p'}$ and ε , using properties of $R_p, R'_p, R_{p'}, R'_{p'}, \delta$, and δ' thoroughly in the next steps:

4a. Note that

$$\begin{aligned} \mathbb{P}_{\mathcal{D}_{n,p,p'}}(S_{p,p'}(n)) &= \mathbb{P}_{\mathcal{D}_{n,p}}(R_p(n)) \cdot \mathbb{P}_{\mathcal{D}_{n,p'}}(R_{p'}(n)) \\ &= (1 - n^{-\omega(n)})(1 - n^{-\omega(n)}) \\ &= 1 - n^{-\omega(n)}, \end{aligned}$$

and similarly for $\mathbb{P}_{\mathcal{D}_{n,p,p'}}(S'_{p,p'}(n))$.

4b. Note that, for any $d \times d' \in S'_{p,p'}$, it holds that $d \in R'_p(n)$ and $d' \in R'_{p'}(n)$. The former implies $\gamma_2 = \lambda(1 - \lambda)(1 + o(1))$ and the latter implies $\gamma'_2 = \lambda'(1 - \lambda')(1 + o(1))$.

4c. By construction, for any $d \times d' \in S_{p,p'}$, it holds that $d \in R_p(n)$ and $d' \in R_{p'}(n)$. Therefore, taking $\varepsilon_{d \times d'} = \delta_d + \delta'_{d'}$, it holds that $|\varepsilon_{d \times d'}| \leq |\delta_d| + |\delta'_{d'}| \leq \delta(n) + \delta'(n) = \varepsilon(n)$, and

$$\begin{aligned} \frac{\mathbb{P}_{\mathcal{D}_{n,p,p'}}(d \times d')}{\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d')} &= \frac{\mathbb{P}_{\mathcal{D}_{n,p}}(d) \cdot \mathbb{P}_{\mathcal{D}_{n,p'}}(d')}{\mathbb{P}_{\mathcal{E}'_{n,p}}(d) \cdot \mathbb{P}_{\mathcal{E}'_{n,p'}}(d')} \\ &= \exp\left\{\frac{1}{4}\left(1 - \frac{\gamma_2^2}{\lambda^2(1 - \lambda)^2}\right)\right\} \cdot \exp\{\delta_d\} \\ &\quad \cdot \exp\left\{\frac{1}{4}\left(1 - \frac{(\gamma'_2)^2}{(\lambda')^2(1 - \lambda')^2}\right)\right\} \cdot \exp\{\delta'_{d'}\} \\ &= \exp\left\{\frac{1}{4}\left(2 - \frac{\gamma_2^2}{\lambda^2(1 - \lambda)^2} - \frac{(\gamma'_2)^2}{(\lambda')^2(1 - \lambda')^2}\right)\right\} \cdot \exp\{\delta_d + \delta'_{d'}\} \\ &= \exp\left\{\frac{1}{4}\left(2 - \frac{\gamma_2^2}{\lambda^2(1 - \lambda)^2} - \frac{(\gamma'_2)^2}{(\lambda')^2(1 - \lambda')^2}\right)\right\} \cdot \exp\{\varepsilon_{d \times d'}\}. \end{aligned}$$

□

Using the above stepwise approximation through the models, we can derive a general-purpose rule for vanishing probabilities of events involving independent $G(n, p)$ random graphs, similar to the one stated in Theorem 2.4.

Theorem 3.2. *Let A_n be a sequence of events in E_n^2 . If p, p' satisfy $\omega(\log n/n) \leq p, p' \leq o(n^{-1/2})$, then $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n) = o(n^{-a})$ implies $\mathbb{P}_{\mathcal{D}_{n,p,p'}}(A_n) = o(n^{-a})$ for any fixed $a > 0$.*

Proof. Before anything, we note that our hypotheses imply that $\omega(1/n) \leq p, p' \leq o(1)$, which we will use several times along the proof. Let $a > 0$ be fixed, and assume $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n) = o(n^{-a})$.

In agreement with the approximation scheme previously presented, we will prove our assertion in four steps, each addressing one of the following statements:

- (1) $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n) = o(n^{-a})$ implies $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) = o(n^{-a})$;
- (2) $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) = o(n^{-a})$ implies $\mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n) = o(n^{-a})$;
- (3) $\mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n) = o(n^{-a})$ implies $\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n) = o(n^{-a})$;
- (4) $\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n) = o(n^{-a})$ implies $\mathbb{P}_{\mathcal{D}_{n,p,p'}}(A_n) = o(n^{-a})$.

Step 1: Assume $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n) = o(n^{-a})$. Theorem 3.1(1) states that

$$\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) = \frac{4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n)}{[1 + (q - p)^{2N}][1 + (q' - p')^{2N}]}.$$

Since $p = \omega(1/n)$, it follows that $2Np \rightarrow \infty$ and $(q - p)^{2N} = (1 - 2Np/2N)^{2N} \rightarrow 0$. By an analogous argument, $(q' - p')^{2N} \rightarrow 0$. Thus, $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) \sim 4\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n)$ and, since $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n) = o(n^{-a})$, it follows that $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) = o(n^{-a})$.

Step 2: Assume $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n) = o(n^{-a})$. We turn to the expression that links $\mathcal{E}_{n,p,p'}$ to $\mathcal{I}_{n,p,p'}$, presented in Theorem 3.1(2).

The normalization constant $V_{n,p}V_{n,p'}$ is the probability that two independent random variables $N(p, pq/2N)$ and $N(p', p'q'/2N)$ assume values in $[0, 1]$. Standardizing these random variables and denoting by $Q(\cdot)$ the Q-function (tail distribution of a standard normal random variable), we have

$$\begin{aligned} V_{n,p} &= Q\left(-\frac{p}{\sqrt{pq/2N}}\right) - Q\left(\frac{q}{\sqrt{pq/2N}}\right) \\ &= Q\left(-\sqrt{\frac{2Np}{q}}\right) - Q\left(\sqrt{\frac{2Nq}{p}}\right) \\ &\rightarrow 1, \end{aligned}$$

where the limit comes from the facts that $2Np/q = \omega(1)$ whenever $p = \omega(1/n)$ and $2Nq/p = \omega(1)$ whenever $p = o(1)$. The same limit applies to $V_{n,p'}$ by the same argument. Thus $1/V_{n,p}V_{n,p'} = \theta(1)$.

For the integral, we will split the domain of integration into several rectangles and deal with them separately. To simplify our notation, we denote our integrand by $g(x, x') = \phi(x; p, \frac{pq}{2N})\phi(x'; p', \frac{p'q'}{2N})\mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n)$. Pick some constant $c > a$, and let $\delta = \delta(n) = c\sqrt{q \log n/Np}$ and $\delta' = \delta'(n) = c\sqrt{q' \log n/Np'}$. Note that $np = \omega(\log n)$ implies $\delta = c\sqrt{pq \log n/Np^2} = o(pq \log n / \log^2 n) = o(p)$. Similarly, $np' = \omega(\log n)$ implies $\delta' = o(p')$.

Since $p = o(q)$ whenever $p = o(1)$, it holds that $\delta < p, q$ for large enough n . For such n , we can safely write

$$\begin{aligned}
& \int_0^1 \int_0^1 \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) dx dx' \\
&= \int_0^1 \int_0^1 g(x, x') dx dx' \\
&= \iint_{(1)} g(x, x') dx dx' + \iint_{(2)} g(x, x') dx dx' + \iint_{(3)} g(x, x') dx dx' \\
&+ \iint_{(4)} g(x, x') dx dx' + \iint_{(5)} g(x, x') dx dx' + \iint_{(6)} g(x, x') dx dx' \\
&+ \iint_{(7)} g(x, x') dx dx' + \iint_{(8)} g(x, x') dx dx' + \iint_{(9)} g(x, x') dx dx',
\end{aligned}$$

where the subdomains of integration are as follows:

$$\begin{aligned}
(1) &= [0, p(1 - \delta)) \times [0, p'(1 - \delta')), \\
(2) &= [0, p(1 - \delta)) \times [p'(1 - \delta'), p'(1 + \delta')], \\
(3) &= [0, p(1 - \delta)) \times (p'(1 + \delta'), 1], \\
(4) &= [p(1 - \delta), p(1 + \delta)] \times [0, p'(1 - \delta')), \\
(5) &= [p(1 - \delta), p(1 + \delta)] \times [p'(1 - \delta'), p'(1 + \delta')], \\
(6) &= [p(1 - \delta), p(1 + \delta)] \times (p'(1 + \delta'), 1], \\
(7) &= (p(1 + \delta), 1] \times [0, p'(1 - \delta')), \\
(8) &= (p(1 + \delta), 1] \times [p'(1 - \delta'), p'(1 + \delta')], \\
(9) &= (p(1 + \delta), 1] \times (p'(1 + \delta'), 1].
\end{aligned}$$

These subdomains are illustrated in Figure 1.

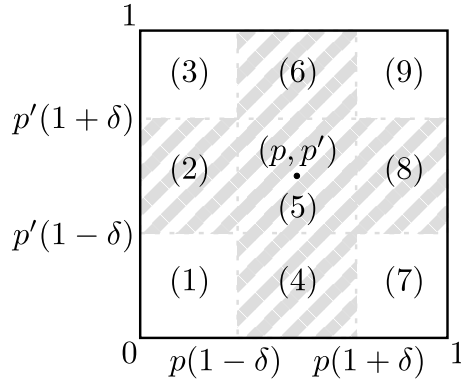


FIGURE 1. Splitting $[0, 1] \times [0, 1]$ into 9 smaller domains of integration. We call domains (1), (3), (7), and (9) the *corner* subdomains, domains (2), (4), (6), and (8) the *side* subdomains, and domain (5) the *central* subdomain.

The integral over the corner subdomain (1) can be bounded easily, by noting that $\mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) \leq 1$:

$$\begin{aligned}
& \int_0^{p(1-\delta)} \int_0^{p(1-\delta')} g(x, x') dx dx' \\
&= \int_0^{p(1-\delta)} \int_0^{p(1-\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) dx dx' \\
&\leq \int_0^{p(1-\delta)} \int_0^{p(1-\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) dx dx' \\
&= \left[\int_0^{p(1-\delta)} \phi(x; p, \frac{pq}{2N}) dx \right] \left[\int_0^{p(1-\delta')} \phi(x'; p', \frac{p'q'}{2N}) dx' \right] \\
&\leq Q \left(\frac{p\delta}{\sqrt{pq/2N}} \right) Q \left(\frac{p'\delta'}{\sqrt{p'q'/2N}} \right) \leq \exp \left\{ -\frac{Np\delta^2}{q} \right\} \exp \left\{ -\frac{Np'(\delta')^2}{q'} \right\} \\
&= \exp\{-c \log n\} \exp\{-c \log n\} = n^{-2c} = o(n^{-a}),
\end{aligned}$$

where the last step comes from the choice of c . Similarly we assert this bound for subdomains (3), (7), and (9), with a few changes in the arguments of the Q-functions.

For the integral on side subdomain (2), we can follow a similar strategy, but a little bit more carefully:

$$\begin{aligned}
& \int_0^{p(1-\delta)} \int_{p(1-\delta')}^{p(1+\delta')} g(x, x') dx dx' \\
&= \int_0^{p(1-\delta)} \int_{p(1-\delta')}^{p(1+\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) dx dx' \\
&\leq \int_0^{p(1-\delta)} \int_{p(1-\delta')}^{p(1+\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) dx dx' \\
&= \left[\int_0^{p(1-\delta)} \phi(x; p, \frac{pq}{2N}) dx \right] \left[\int_{p(1-\delta')}^{p(1+\delta')} \phi(x'; p', \frac{p'q'}{2N}) dx' \right] \\
&\leq Q \left(\frac{p\delta}{\sqrt{pq/2N}} \right) \cdot 1 \leq \exp \left\{ -\frac{Np\delta^2}{q} \right\} \\
&= \exp\{-c \log n\} = n^{-c} = o(n^{-a}).
\end{aligned}$$

Again, the same bound holds for subdomains (4), (6), and (8), with minor changes in the arguments of the Q-functions.

For the integral on center subdomain (5), some comments are appropriate. First, note that, since $\delta = o(p)$, for any $x = x(n) \in [p(1-\delta), p(1+\delta)]$ it is true that $x = p(1+o(p))$ and, therefore, x has the same asymptotics as p — namely, $o(n^{-1/2}) \leq x \leq \omega(\log n/n)$. Similarly, for any $x' = x'(n) \in [p'(1-\delta'), p'(1+\delta')]$, it holds that $x' = o(n^{-1/2})$ and $x' = \omega(\log n/n)$.

Also, for any fixed n , $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n)$ is a continuous function of p and p' . This comes from the result of Theorem 3.1(1) and the fact that $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n) = \sum_{d \times d' \in A_n} \mathbb{P}_{\mathcal{B}_{n,p,p'}}(d \times d')$. Since the probability of each such $d \times d'$ under the measure $\mathbb{P}_{\mathcal{B}_{n,p,p'}}$ is a continuous function of p and p' (product of powers of p , p' , $1-p$ and $1-p'$ and some constants in both p and p'), and the sum

of these functions has a finite number of terms, continuity of $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n)$ with respect to p and p' follows. Since, by Theorem 3.1(1), $\mathbb{P}_{\mathcal{E}_{n,p,p'}}(A_n)$ is the product between $\mathbb{P}_{\mathcal{B}_{n,p,p'}}(A_n)$ and a continuous function of p and p' , continuity of the former with respect to p and p' also follows.

As a consequence of these results, for $(x, x') \in [p(1-\delta), p(1+\delta)] \times [p'(1-\delta'), p'(1+\delta')]$, the function $\mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n)$, being a continuous function over this compact set, will attain a maximum value for some argument $(y, y')(n)$ in this set. Such (y, y') will, forcefully, satisfy $y, y' = o(n^{-1/2})$ and $y, y' = \omega(\log n/n)$, which means, by our conclusion from the previous step, that $\mathbb{P}_{\mathcal{E}_{n,y,y'}}(A_n) = o(n^{-a})$.

That being said, we can assert that

$$\begin{aligned}
& \int_{p(1-\delta)}^{p(1+\delta)} \int_{p'(1-\delta')}^{p'(1+\delta')} g(x, x') dx dx' \\
&= \int_{p(1-\delta)}^{p(1+\delta)} \int_{p'(1-\delta')}^{p'(1+\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) dx dx' \\
&\leq \int_{p(1-\delta)}^{p(1+\delta)} \int_{p'(1-\delta')}^{p'(1+\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \cdot \\
&\quad \left[\max_{(x,x') \in [p(1-\delta), p(1+\delta)] \times [p'(1-\delta'), p'(1+\delta')]} \mathbb{P}_{\mathcal{E}_{n,x,x'}}(A_n) \right] dx dx' \\
&= \int_{p(1-\delta)}^{p(1+\delta)} \int_{p'(1-\delta')}^{p'(1+\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) \mathbb{P}_{\mathcal{E}_{n,y,y'}}(A_n) dx dx' \\
&= \mathbb{P}_{\mathcal{E}_{n,y,y'}}(A_n) \int_{p(1-\delta)}^{p(1+\delta)} \int_{p'(1-\delta')}^{p'(1+\delta')} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) dx dx' \\
&\leq \mathbb{P}_{\mathcal{E}_{n,y,y'}}(A_n) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x; p, \frac{pq}{2N}) \phi(x'; p', \frac{p'q'}{2N}) dx dx' \\
&= o(n^{-a}) \cdot 1 = o(n^{-a}).
\end{aligned}$$

Thus, we conclude that

$$\mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n) = \theta(1) \cdot (9 \cdot o(n^{-a})) = o(n^{-a}).$$

Step 3: Assume $\mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n) = o(n^{-a})$. We begin by recalling that

$$\left(\frac{1}{2}M(S_1), \frac{1}{2}M(S_2) \right) \stackrel{d}{\sim} \text{Bin}(N, p) \otimes \text{Bin}(N, p') \text{ under } \mathbb{P}_{\mathcal{E}'_{n,p,p'}}.$$

Define the event $N_n = \{|M(S_1) - 2Np| < 2Np \cdot \varepsilon, |M(S_2) - 2Np'| < 2Np' \cdot \varepsilon'\}$, with $\varepsilon = (2Np)^{-5/12}$ and $\varepsilon' = (2Np')^{-5/12}$. By the Chernoff bound, we have

$$\begin{aligned}
& \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(|M(S_1) - 2Np| \geq 2Np \cdot \varepsilon) \leq 2e^{-2Np\varepsilon^2/6} = 2e^{-\frac{1}{6}(2Np)^{1/6}}, \\
& \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(|M(S_2) - 2Np'| \geq 2Np' \cdot \varepsilon') \leq 2e^{-2Np'(\varepsilon')^2/6} = 2e^{-\frac{1}{6}(2Np')^{1/6}}.
\end{aligned}$$

Thus, by the union bound, $\mathbb{P}_{\mathcal{E}'_{n,p,p}}(\overline{N_n}) \leq 2(e^{-\frac{1}{6}(2Np)^{1/6}} + e^{-\frac{1}{6}(2Np')^{1/6}})$.

Now, by the definition of the event N_n , it holds that, in this event:

$$\begin{aligned} |M(S_1) - 2Np| &< 2Np \cdot \varepsilon = (2Np)(2Np)^{-5/12} \frac{\sqrt{4Npq}}{\sqrt{4Npq}} \\ &= \left[\frac{(2Np)^{1/12}}{\sqrt{2q}} \right] \sqrt{4Npq}, \\ |M(S_2) - 2Np'| &< 2Np' \cdot \varepsilon' = (2Np')(2Np')^{-5/12} \frac{\sqrt{4Np'q'}}{\sqrt{4Np'q'}} \\ &= \left[\frac{(2Np')^{1/12}}{\sqrt{2q'}} \right] \sqrt{4Np'q'}. \end{aligned}$$

The same inequalities hold in the event $A_n \cap N_n \subseteq N_n$. Now, note that $(2Np)^{1/12}/\sqrt{2q} = o(\sqrt[5]{2Npq})$ and $(2Np')^{1/12}/\sqrt{2q'} = o(\sqrt[5]{2Np'q'})$, which allows us to relate the probability of $A_n \cap N_n$ under measures $\mathbb{P}_{\mathcal{E}'_{n,p,p'}}$ and $\mathbb{P}_{\mathcal{I}_{n,p,p'}}$. We choose $y = \max\{(2Np)^{1/12}/\sqrt{2q}, (2Np')^{1/12}/\sqrt{2q'}\}$; this choice of y and $q = \theta(1)$ imply $(1+|y|^3)/\sqrt{pqN} = o((Np)^{-1/2}) + o(n^{3/8})/\omega(n^{1/2}) = o(1)$ and $(1+|y|^3)/\sqrt{p'q'N} = o((Np')^{-1/2}) + o(n^{3/8})/\omega(n^{1/2}) = o(1)$. From these facts, using Theorem 3.1(3), it follows that

$$\begin{aligned} \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n) &= \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n \cap \overline{N_n}) + \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n \cap N_n) \\ &\leq \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(\overline{N_n}) + \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n \cap N_n) \\ &\leq 2(e^{-\frac{1}{6}(2Np)^{1/6}} + e^{-\frac{1}{6}(2Np')^{1/6}}) \\ &\quad + \mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n \cap \overline{N_n}) \left(1 + O\left(\frac{1+|y|^3}{\sqrt{pqN}}\right) + O\left(\frac{1+|y|^3}{\sqrt{p'q'N}}\right) \right)^{-1} \\ &\leq e^{-\omega(n)} + \mathbb{P}_{\mathcal{I}_{n,p,p'}}(A_n)(1 + o(1) + o(1))^{-1} \\ &= o(n^{-a}) + o(n^{-a})(\theta(1))^{-1} = o(n^{-a}). \end{aligned}$$

Step 4: Assume $\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n) = o(n^{-a})$. Let the sets $S_{p,p'}(n), S'_{p,p'}(n)$ and the real function $\varepsilon(n)$ be as in Theorem 3.1(4) (note that our hypotheses about p, p' imply the hypotheses of this theorem are satisfied), and define the set $T_{p,p'}(n) = S_{p,p'}(n) \cap S'_{p,p'}(n)$. Then the following facts hold:

- (1) $\mathbb{P}_{\mathcal{D}_{n,p,p'}}(T_{p,p'}(n)) = 1 - n^{-\omega(n)}$ (as, by union bound, $\mathbb{P}_{\mathcal{D}_{n,p,p'}}(\overline{T_{p,p'}(n)}) \leq \mathbb{P}_{\mathcal{D}_{n,p,p'}}(\overline{S_{p,p'}(n)}) + \mathbb{P}_{\mathcal{D}_{n,p,p'}}(\overline{S'_{p,p'}(n)}) = 2n^{-\omega(n)} = n^{-\omega(n)}$);
- (2) for every $d \times d' \in T_{p,p'}(n)$, there is some $\varepsilon_{d \times d'}$ such that $|\varepsilon_{d \times d'}| \leq \varepsilon(n)$ and

$$\frac{\mathbb{P}_{\mathcal{D}_{n,p,p'}}(d \times d')}{\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d')} = \exp \left\{ \frac{1}{4} \left(2 - \frac{\gamma_2^2}{\lambda^2(1-\lambda)^2} - \frac{(\gamma'_2)^2}{(\lambda')^2(1-\lambda')^2} \right) \right\} \cdot \exp\{\varepsilon_{d \times d'}\};$$

- (3) in $T_{p,p'}(n)$, $\gamma_2 = \lambda(1-\lambda)(1+o(1))$ and $\gamma'_2 = \lambda'(1-\lambda')(1+o(1))$;

Using these facts, it follows that:

$$\begin{aligned}
\mathbb{P}_{\mathcal{D}_{n,p,p'}}(A_n) &= \mathbb{P}_{\mathcal{D}_{n,p,p'}}(A_n \cap \overline{T_{p,p'}(n)}) + \mathbb{P}_{\mathcal{D}_{n,p,p'}}(A_n \cap T_{p,p'}(n)) \\
&\leq \mathbb{P}_{\mathcal{D}_{n,p,p'}}(\overline{T_{p,p'}(n)}) + \sum_{d \times d' \in A_n \cap T_{p,p'}(n)} \mathbb{P}_{\mathcal{D}_{n,p,p'}}(d \times d') \\
&= n^{-\omega(n)} + \sum_{d \times d' \in A_n \cap T_{p,p'}(n)} \left[\mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d') \cdot \right. \\
&\quad \left. \exp \left\{ \frac{1}{4} \left(2 - \frac{\gamma_2^2}{\lambda^2(1-\lambda)^2} - \frac{(\gamma'_2)^2}{(\lambda')^2(1-\lambda')^2} \right) \right\} \cdot \exp\{\varepsilon_{d \times d'}\} \right] \\
&= n^{-\omega(n)} + \left[\sum_{d \times d' \in A_n \cap T_{p,p'}(n)} \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(d \times d') \right] \cdot \\
&\quad \max_{d \times d' \in A_n \cap T_{p,p'}(n)} \exp \left\{ \frac{1}{4} \left(2 - \frac{\gamma_2^2}{\lambda^2(1-\lambda)^2} - \frac{(\gamma'_2)^2}{(\lambda')^2(1-\lambda')^2} \right) \right\} \cdot \\
&\quad \max_{d \times d' \in A_n \cap T_{p,p'}(n)} \exp\{\varepsilon_{d \times d'}\} \\
&\leq n^{-\omega(n)} + \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n \cap T_{p,p'}(n)) \cdot \\
&\quad \exp \left\{ \frac{1}{4} (2 - (1 + o(1))^2 - (1 + o(1))^2) \right\} \cdot \exp\{\varepsilon(n)\} \\
&\leq o(n^{-a}) + \mathbb{P}_{\mathcal{E}'_{n,p,p'}}(A_n) \cdot \exp\{o(1)\} \cdot \exp\{o(1)\} \\
&= o(n^{-a}) + o(n^{-a}) \cdot \theta(1) \cdot \theta(1) = o(n^{-a}).
\end{aligned}$$

□

4. EXAMPLE APPLICATION

The results from the previous section establish an approximation scheme between the degree sequences of $G(n, p)$ random graphs and sequences of independent binomial random variables. As such, it allows us to determine properties of random graphs using a much simpler and more well-studied object. Intuitively, if a graph property is related to some feature of its degree sequence, one can take this feature as a proxy for the original property, analyze it assuming the degrees are independent (that is, under the $B_{n,p,p'}$ model), and use the framework to carry over the findings.

As an example application, consider the problem of *graph isomorphism*: given two graphs G_1 and G_2 , we would like to determine whether or not they are isomorphic, that is, whether there is an edge-preserving mapping between their vertex sets. While this is an interesting problem, and vastly explored in graph theory from a deterministic point of view, it can also be studied in probabilistic settings, such as that in which G_1 and G_2 are drawn from known random graph models. Most of the work in this kind of setting follows an algorithmic approach, i.e., an algorithm is sought which correctly asserts a.a.s. whether G_1 and G_2 are isomorphic. Note that the asymptotic correctness of the algorithm will, in general, depend on the random graph model of choice, including the regimes of its parameters [2]. Moreover, the use of *canonical labeling algorithms* [4] is often preferred, which reduces checking whether two graphs are isomorphic to verifying whether their canonical forms are precisely equal [2][3][10][12] (this can be done in quadratic time over the size of the graphs' vertex sets).

Here, by contrast, we follow a structural approach to the problem, i.e., we would like to determine whether two graphs G_1 and G_2 are isomorphic or not isomorphic

a.a.s. Problems of this nature require a mathematical solution (i.e., a proof), rather than an algorithmic solution². In our example, we proceed by assuming that G_1 and G_2 are independent $G(n, p)$ random graphs. In this case, the following result holds:

Theorem 4.1. *Let G, G' be two independent random graphs, distributed as $G(n, p)$ and $G(n, p')$ respectively, for some $p, p' \in (0, 1)$. If $\omega(\log n/n) \leq p, p' \leq o(n^{-1/2})$, then $\mathbb{P}_{\mathcal{D}_{n,p,p'}}[G, G' \text{ are isomorphic}] = o(n^{-1/2})$.*

To prove this result, we will use an auxiliary graph-theoretic proposition. Denote by $d_G(v)$ the degree of vertex v in graph G . For an arbitrary Borel set B on the real line, define $F_B(G) = |\{v \in V(G) : d_G(v) \in B\}|$, that is, $F_B(G)$ counts the number of vertices in G with degrees in B . At the expense of a slight abuse of notation, for any finite sequence d of length $|d|$, denote $F_B(d) = \{i \in [|d|] : d[i] \in B\}$, where $d[i]$ is the i -th component of d . Note that, if d is the degree sequence of graph G , then $F_B(G) = F_B(d)$.

Proposition 4.2. *If G, G' are isomorphic, then for every Borel set B on the real line, $F_B(G) = F_B(G')$.*

Proof. Let $f : V(G) \rightarrow V(G')$ be an isomorphism between G and G' (since G and G' are isomorphic, there is at least one such f). f is, by definition, bijective. Also, since f is edge-preserving, f is also degree-preserving, that is, $d_G(v) = d_{G'}(f(v))$ for any $v \in V(G)$. Using these facts, for every Borel set B , we have

$$\begin{aligned} F_B(G) &= |\{v \in V(G) : d_G(v) \in B\}| \\ &= |\{v \in V(G) : d_{G'}(f(v)) \in B\}| \\ &= |\{v' \in V(G') : d_{G'}(v') \in B\}| = F_B(G'). \end{aligned}$$

□

We can now proceed to the proof of Theorem 4.1:

Proof. Let $B_n = [\lfloor \min\{(n-1)p, (n-1)p'\} \rfloor, \infty)$. Proposition 4.2 implies

$$\begin{aligned} \mathbb{P}_{\mathcal{D}_{n,p,p'}}[G, G' \text{ are isomorphic}] &\leq \mathbb{P}_{\mathcal{D}_{n,p,p'}}[F_{B_n}(G) = F_{B_n}(G')] \\ &= \mathbb{P}_{\mathcal{D}_{n,p,p'}}[F_{B_n}(d) = F_{B_n}(d')], \end{aligned}$$

where d and d' are the degree sequences of G and G' , respectively. We will show that the right-hand side of the inequality is $o(n^{-1/2})$. By Theorem 3.2, it is enough to show that $\mathbb{P}_{\mathcal{B}_{n,p,p'}}[F_{B_n}(d) = F_{B_n}(d')] = o(n^{-1/2})$.

Now, note that, in the $\mathcal{B}_{n,p,p'}$ model, each element of d belongs to B_n independently of the others. Furthermore, each such element is a $\text{Bin}(n-1, p)$ random variable and, therefore, it belongs to B_n with probability $\alpha_n > 1/2$ (since the median of $\text{Bin}(n-1, p)$ is at least $(n-1)p$). This implies $F_{B_n}(d)$ is the sum of n independent $\text{Be}(\alpha_n)$ random variables, hence it has distribution $\text{Bin}(n, \alpha_n)$.

²In particular, in a regime where two random graph instances are isomorphic a.a.s., the trivial algorithm that always outputs “YES” will be correct a.a.s.

Moving on, let $b(k; n, p)$ be the mass function of a $\text{Bin}(n, p)$ random variable. Since $\alpha_n = \omega(1/n)$, it holds that $\max_k b(k; n, \alpha_n) = o(n^{-1/2})$ [6]. This implies

$$\begin{aligned} \mathbb{P}_{\mathcal{B}_{n,p,p'}}[F_{B_n}(d) = F_{B_n}(d')] &= \sum_k \mathbb{P}_{\mathcal{B}_{n,p,p'}}[F_{B_n}(d) = k, F_{B_n}(d') = k] \\ &= \sum_k b(k; n, \alpha_n) \cdot b(k; n, \alpha'_n) \\ &\leq \sum_k \left[\max_{k'} b(k'; n, \alpha_n) \right] \cdot b(k; n, \alpha'_n) \\ &= \max_{k'} b(k'; n, \alpha_n) \cdot \sum_k b(k; n, \alpha'_n) \\ &= o(n^{-1/2}) \cdot 1 = o(n^{-1/2}). \end{aligned}$$

□

Note that this proof could not be easily carried out directly on the degree sequences of G and G' , since the degrees of individual vertices are not independent random variables, which makes it hard in general to determine the distribution of F_{B_n} . However, the result can be alternatively proven without the approximation scheme and for a wider range of p, p' . One way to achieve this is by noting that G, G' are isomorphic only if $M = M'$, and that $M/2, M'/2$ are distributed as binomial random variables, which converge in probability to normal random variables with deviation of at most $o(n^{-1/2})$ by the Berry-Esseen theorem.

Our choice to prove the simpler statement of Theorem 4.1 has aimed not only to highlight an application of our method but also to present a clean and simple proof that can be further modified to show a stronger result:

If $\omega(\log n/n) \leq p, p' \leq o(n^{-1/2})$, then $\mathbb{P}_{\mathcal{D}_{n,p,p'}}[G, G' \text{ are isomorphic}] = o(n^{-a})$ for any $a > 0$.

The context in which we require such stronger statement, as well as its detailed proof, are the subject of ongoing work.

5. FINAL REMARKS

In this paper, we have considered the degree sequences of two independent Erdős-Rényi random graphs and an approximation model in which such degrees are considered to be independent. We have formally shown that any sequence of events in the approximation model with probability smaller than a power law will have this inequality carried over to the original degree-sequence model.

The techniques used in our extension from the original, single-graph framework by McKay and Wormald can be applied to the more general scenario consisting of $k = k(n)$ independent random graphs, rather than only $k = 2$ random graphs, and yield similar asymptotic expressions for bounded k , provided the parameters of the generalized model satisfy analogous assumptions. As an example, letting the i -th graph be distributed as $G(n, p_i)$, and $q_i = 1 - p_i$, the proportionality factor in Theorem 3.1(1) generalizes to

$$\frac{2^k}{\prod_{i \leq k} [(1 + (q_i - p_i)^{2N})]}.$$

We believe the same kind of asymptotic equivalence can also be achieved if k grows slowly (possibly any $k = o(\log n)$), but we leave this matter open for future study.

REFERENCES

- [1] ALON, N. AND SPENCER, J. H. (1992). *The Probabilistic Method*. Wiley, New York.
- [2] BABAI, L., ERDŐS, P. AND SELKOW, S. M. (1980). Random graph isomorphism. *SIAM Journal on Computing* **9**, 628–635.
- [3] BABAI, L. AND KUČERA, L. (1979). Canonical labelling of graphs in linear average time. In *IEEE FOCS*. pp. 39–46.
- [4] BABAI, L. AND LUKS, E. M. (1983). Canonical labeling of graphs. In *ACM STOC*. pp. 171–183.
- [5] BOLLOBÁS, B. (2001). *Random Graphs* second ed. Cambridge University Press, Cambridge, UK.
- [6] CZAJKA, T. AND PANDURANGAN, G. (2008). Improved random graph isomorphism. *Journal of Discrete Algorithms* **6**, 85–92.
- [7] ERDŐS, P. AND RÉNYI, A. (1959). On random graphs I. *Publicationes Mathematicae (Debrecen)* **6**, 290–297.
- [8] FOGGIA, P. AND VENTO, M. (2012). Graph matching techniques for computer vision. *Graph-Based Methods in Computer Vision: Developments and Applications* 1–41.
- [9] GILBERT, E. N. (1959). Random graphs. *The Annals of Mathematical Statistics* **30**, 1141–1144.
- [10] KARP, R. M. (1979). Probabilistic analysis of a canonical numbering algorithm for graphs. In *AMS Symposium in Pure Mathematics*. vol. 34. pp. 365–378.
- [11] KOSTOCHKA, A. V. AND WEST, D. B. (2006). Chvátal’s condition cannot hold for both a graph and its complement. *Discussiones Mathematicae Graph Theory* **26**, 73–76.
- [12] LIPTON, R. J. (1978). *The beacon set approach to graph isomorphism*. Research Report 135, Department of Computer Science, Yale University.
- [13] MCKAY, B. D. AND WORMALD, N. C. (1997). The degree sequence of a random graph. I. The models. *Random Structures and Algorithms* **11**, 97–117.
- [14] PEDARSANI, P. AND GROSSGLAUSER, M. (2011). On the privacy of anonymized networks. In *ACM KDD*. pp. 1235–1243.
- [15] RÁTH, B. (2009). Mean field frozen percolation. *Journal of Statistical Physics* **137**, 459–499.
- [16] SKERMAN, F. (2010). Degree sequences of random bipartite graphs. *PhD thesis*. The Australian National University.